# APPROXIMATING SOLUTIONS OF EQUATIONS BY COMBINING NEWTON-LIKE METHODS

## IOANNIS K. ARGYROS

ABSTRACT. In cases sufficient conditions for the semilocal convergence of Newton-like methods are violated, we start with a modified Newton-like method (whose weaker convergence conditions hold) until we stop at a certain finite step. Then using as a starting guess the point found above we show convergence of the Newton-like method to a locally unique solution of a nonlinear operator equation in a Banach space setting. A numerical example is also provided.

### 1. Introduction

In this study we are concerned with the problem of approximating a locally unique solution  $x^*$  of the nonlinear equation

$$(1) F(x) = 0,$$

where F is a Fréchet-differentiable operator defined on a convex subset D of a Banach space X with values in a Banach space Y.

The most popular methods for generating sequences approximating  $x^*$  are undoubtedly Newton-like methods of the form

(2) 
$$y_{n+1} = y_n - A(y_n)^{-1} F(y_n), \quad (y \in D), \quad (n \ge 0)$$

and the corresponding modified Newton-like methods given by

(3) 
$$x_{n+1} = x_n - A(x_0)^{-1} F(x_n), \quad (x_0 \in D), (n \ge 0).$$

Here  $A(x) \in L(X,Y)$   $(x \in D)$  the space of bounded linear operators from X into Y. The most popular choice for A with is an approximation of F'(x) is usually A(x) = F'(x)  $(x \in D)$ . In this case (2) becomes Newton's method and (3) the corresponding Modified Newton method.

Received by the editors October 27, 2006.

<sup>2000</sup> Mathematics Subject Classification. 65H10, 65G99, 47H10, 49M15.

Key words and phrases. Newton-like method, modified Newton-like method, Banach space, semilocal convergence, Fréchet-derivative, Newton-Kantorovich-type hypotheses.

Local as well as semilocal convergence theorems for methods (2) and (3) have been given by many authors under different assumptions. A survey of such results can be found in [1]–[7] and the references there. Method (2) is faster but more expensive than (3). Moreover the sufficient Newton-Kantorovich type conditions for the semilocal convergence of method (2) are stronger (see (9)) than method's (3) (see (20) or (68)). Here we take advantage of these observations in case (9) is violated for some starting guess  $y_0$  but (20) or (68) hold for  $x_0 = y_0$ . Using formula (60) we can always find a finite integer N such that for  $y_0 = x_N$  conditions (9) holds true. This way after using slower method (3) for N steps we continue with faster method (2) to obtain convergence. A numerical example is also provided where our technique is demonstrated.

### 2. Semilocal Convergence Analysis

We state the well-Known semilocal convergence result for Newton-like method (2), [7], [2].

**Theorem 2.1.** Let  $F: D \subset X \to Y$  be Fréchet differentiable and let  $A(x) \in L(X,Y)$  be an approximation of F'(x). Assume there exist  $y_0 \in D$ , the inverse  $A(y_0)^{-1}$  of  $A(x_0)$  and constants  $\eta, K > 0$ ,  $M, L, \mu, l \geq 0$  such that for all  $x, y \in D$  the following conditions hold:

$$||A(y_0)^{-1}F(y_0)|| \le \mu$$

(5) 
$$||A(y_0)^{-1}[F'(x) - F'(y)]|| \le K ||x - y||,$$

(6) 
$$||A(y_0)^{-1}[F'(x)|| - A(x)] < M ||x - y_0|| + \mu,$$

(7) 
$$||A(y_0)^{-1}[A(x) - A(y_0)]|| \le L ||x - y_0|| + l,$$

$$(8) b = \mu + l < 1,$$

$$(9) h = \sigma \eta \le \frac{1}{2} (1 - b)^2$$

where,

(10) 
$$\sigma = \max\{K, M + L\},\$$

and

(11) 
$$\bar{U}(y_0, t^*) = \{x \in X \mid ||x - y_0|| \le t^*\} \subset D,$$

where t\*is the smallest zero of scalar function

(12) 
$$f(t) = \frac{\sigma}{2}t^2 - (1-b)t + \eta$$

given by

(13) 
$$t^* = \frac{1 - b - \sqrt{(1 - b)^2 - 2h}}{\sigma},$$

with the larger zero t\*\* given by

(14) 
$$t^{**} = \frac{1 - b + \sqrt{(1 - b)^2 - 2h}}{\sigma}.$$

Then,

(a) Scalar sequence  $\{t_n\}(n \geq 0)$  generated by

(15) 
$$t_{n+1} = t_n + \frac{f(t_n)}{1 - l - Lt_n}, \quad t_0 = 0$$

is nondecreasing and converges to t\*.

(b) Sequence  $\{y_n\}$  generated by Newton-like method (2) is well defined, remains in  $\bar{U}(y_0, t^*)$  for all  $n \geq 0$ , and converges to a unique solution  $x^*$  in  $\bar{U}(y_0, t^*) \cap D$  if  $h = \frac{1}{2}(1-b)^2$  or in  $U(y_0, t^{**}) \cap D$  if  $h < \frac{1}{2}(1-b)^2$ . Moreover the following estimates hold for all  $n \geq 0$ :

(16) 
$$||y_{n+1} - y_n|| \le t_{n+1} - t_n,$$
 and

$$||y_n - x^*|| \le t^* - t_n.$$

We can show the main semilocal convergence theorem for the modified Newtonlike method (3).

**Theorem 2.2.** Let  $F: D \subset X \to Y$  be a Fréchet differentiable operator. For  $x_0 \in D$  let  $A(x_0) \in L(X,Y)$  such that  $A(x_0)^{-1} \in L(Y,X)$ , and (4) hold (with  $y_0 = x_0$ ). Assume there exist constants  $M_0 \geq 0$ ,  $\mu_0 \in [0,1)$  such that together with (5) the following conditions hold:

(18) 
$$||A(x_0)^{-1}[F'(x) - A(x_0)]|| \le M_0 ||x - x_0|| + \mu_0, \quad \text{for all } x \in D,$$

$$(19) M_0 \leq K,$$

(20) 
$$h_A = K\eta \le \frac{1}{2}(1 - \mu_o)^2$$

and

$$(21) \bar{U}(x_0, r^*) \subseteq D,$$

where  $r^*$  is the smallest of the two positive zeros of function

(22) 
$$g(r) = \frac{K}{2}r^2 - (1 - \mu_0)r + \eta,$$

given by

(23) 
$$r^* = \frac{1 - \mu_0 - \sqrt{(1 - \mu_0)^2 - 2h_A}}{K}.$$

Note that the largest root  $r^{**}$  of g is given by

(24) 
$$r^{**} = \frac{1 - \mu_0 + \sqrt{(1 - \mu_0)^2 - 2h_A}}{K}.$$

Then,

(a) Scalar sequence  $\{r_n\}$   $(n \ge 0)$  generated by

(25) 
$$r_{n+1} = r_n + g(r_n) \quad (r_0 = 0), \quad (n \ge 0)$$

is nondecreasing, and converges to  $r^*$ .

(b) Sequence  $\{x_n\}$  generated by modified Newton-like method (3) remains in  $\bar{U}(x_0,r^*)$  for all  $n\geq 0$ , and converges to a unique solution  $x^*$  of equation F(x) = 0 in  $\bar{U}(x_0, r^*)$ . Moreover the following estimates hold for all  $n \geq 0$ :

$$||x_{n+1} - x_n|| \le r_{n+1} - r_n,$$

and

$$||x_n - x^*|| \le r^* - r_n.$$

Furthermore, if there exists  $R \geq r^*$  such that

(28) 
$$\frac{M_0}{2}(r^* + R) + \mu_0 \le 1.$$

then the solution  $x^*$  is unique in  $U(x_0, R)$ 

*Proof.* (a) It follows by (20) that function g has two positive zeros  $r^*$  and  $r^{**}$  with  $r^* \leq r^{**}$ . Simple induction on  $n \geq 0$  shows that  $r_n \leq r_{n+1} \leq r^*$  for all  $n \geq 0$ . That is, sequence  $\{r_n\}$  is nondecreasing, and bounded above by  $r^*$ . In view of (25), we get  $r^* = \lim_{n \to \infty} r_n$ . (b) We shall show:

$$||x_{k+1} - x_k|| \le r_{k+1} - r_k,$$

and

(30) 
$$\bar{U}(x_{k+1}, r^* - r_{k+1}) \subset \bar{U}(x_k, r^* - r_k)$$

hold for all  $k \geq 0$ .

For every  $z \in \bar{U}(x_1, r^* - r_1)$ ,

$$||z - x_0|| \le ||z - x_1|| + ||x_1 - x_0|| < r^* - r_1 + r_1 = r^* - r_0$$

implies  $z \in \bar{U}(x_0, r^* - r_0)$ . Since also

$$||x_1 - x_0|| = ||A(x_0)^{-1}F(x_0)|| \le \eta = r_1 - r_0,$$

(29) and (30) hold for k = 0. Given they hold for n = 0, 1, ..., k, then

$$||x_{k+1} - x_0|| \le \sum_{i=1}^{k+1} ||x_i - x_{i-1}|| \le \sum_{i=1}^{k+1} (r_i - r_{i-1}) = r_{k+1} - r_0 = r_{k+1},$$

and

$$||x_k + t(x_{k+1} - x_k) - x_0|| \le r_k + t(r_{k+1} - r_k) \le r^*, \ t \in [0, 1].$$

Using (3) we obtain the approximation

(31) 
$$F(x_{k+1}) = F(x_{k+1}) - F(x_k) - A(x_0)(x_{k+1} - x_k)$$

$$= \int_0^1 [F'(x_k + t(x_{k+1} - x_k)) - A(x_0)](x_{k+1} - x_k)dt$$

$$= \int_0^1 [F'(x_k + t(x_{k+1} - x_k)) - F'(x_k)](x_{k+1} - x_k)dt$$

$$+ [F'(x_k) - A(x_0)](x_{k+1} - x_k).$$

In view of (3), (5), (18), (19), (25), (31), and the induction hypotheses, we obtain in turn

$$(32) ||x_{k+2} - x_{k+1}|| = ||A(x_0)^{-1} F(x_{k+1})||$$

$$\leq \frac{K}{2} ||x_{k+1} - x_k||^2 + [M_0 ||x_k - x_0|| + \mu_0] ||x_{k+1} - x_k||$$

$$\leq \frac{K}{2} (r_{k+1} - r_k)^2 + [M_0 r_k + \mu_0] (r_{k+1} - r_k)$$

$$= \frac{K}{2} (t_{k+1} - t_k)^2 + M_0 t_k (t_{n+1} - t_n) + \mu_0 (t_{k+1} - t_k)$$

$$- (t_{k+1} - t_k) + g(t_k)$$

$$= g(t_{k+1}) + (M_0 - k) t_k t_{k+1} - (M_0 - k) t_k^2$$

$$= g(t_{k+1}) + t_k (M_0 - K) (t_{k+1} - t_k) \leq g(t_{k+1}) = t_{k+2} - t_{k+1},$$

which shows (29) for all  $k \geq 0$ .

Thus for every  $z \in \bar{U}$   $(x_{k+2}, r^* - r_{k+2})$  we have

$$||z - x_{k+1}|| \le ||z - z_{k+2}|| + ||x_{k+2} - x_{k+1}|| \le r^* - r_{k+2} + r_{k+2} - r_{k+1}$$
  
=  $r^* - r_{k+1}$ ,

which implies

$$(33) z \in \bar{U}(x_{k+1}, r^* - r_{k+1}).$$

The induction for (29) and (30) is now completed.

Part (a) implies that sequence  $\{t_n\}$  is Cauchy. In view of (32) and (33) it follows that  $\{x_n\}$  is a Cauchy sequence in a Banach space X, and as such it converges to some  $x^* \in \bar{U}(x_0, r^*)$  (since  $\bar{U}(x_0, r^*)$  is a closed set). By letting  $k \to \infty$  in (32) we obtain  $F(x^*) = 0$ . Estimate (27) follows from (26) by using standard majorization techiques [2], [5], [6]. To show uniqueness first in  $\bar{U}(x_0, r^*)$ , let  $y^*$  be a solution of equation (1) in  $\bar{U}(x_0, r^*)$ . Using (2), (18) and (23) we obtain in turn

$$x_{n+1} - y^* = x_n - y^* - A(x_0)^{-1} F(x_n)$$

$$= -A(x_0)^{-1} \int_0^1 [F'(y^* + t(x_n - y^*)) - A(x_0)](x_n - y^*)$$

and

$$||x_{n+1} - y^*|| \le \int_0^1 [M_0 ||x^* + t(x_n - y^*) - x_0|| + \mu_0] dt ||x_n - y^*||$$

$$\le \int_0^1 M_0 [t ||x_n - x_0|| + (1 - t) ||y^* - x_0|| + \mu_0 dt ||x_n - y^*||$$

$$\le \int_0^1 M_0 [tr^* + (1 - t)r^* + \mu_0] dt ||x_n - y^*||$$

$$\le (M_0 r^* + \mu_0) ||x_n - y^*||$$

$$< ||x_n - y^*||,$$

since

$$(35) q = M_0 r^* + \mu_0 \in [0, 1)$$

by the choice of  $r^*$ .

It follows from (34) that  $\lim_{n\to\infty} x_n = y^*$ . But we showed  $\lim_{n\to\infty} x_n = x^*$ . Hence, we deduce

$$(36) x^* = y^*.$$

Finally let  $y^* \in U(x_0, R)$  be a solution of equation F(x) = 0, where R satisfies (28). As in (34) we obtain

(37) 
$$||x_{n+1} - y^*|| < \left[\frac{M_0}{2}(r^* + R) + \mu_0\right] ||x_n - y^*|| \le ||x_n - y^*||,$$

which leads to (36). That completes the proof of the Theorem.

**Remark 2.3.** (a) If  $R = r^{**}$  then condition (28) can be replaced by

$$(38) M_0 \le 1.$$

- (b) It follows from (34) that the order of convergence of method (3) is at least linear with geometric ratio  $q_0$  given by (35).
- (c) Under hypotheses (6) and (7)

(39) 
$$||A(x_0)^{-1}[F'(x) - A(x_0)]|| \le ||A(x_0)^{-1}[F'(x) - A(x)]||$$

$$+ ||A(x_0)^{-1}[A(x) - A(x_0)]||$$

$$\le (M+L)||x - x_0|| + \mu + l.$$

In view of (18) we get

$$(40) M_0 < M + L,$$

and

Therefore, we deduce

(42) 
$$h \le \frac{1}{2}(1-b)^2 \Longrightarrow h_A \le \frac{1}{2}(1-\mu_0)^2,$$

but not viceversa unless if  $\sigma = k$ ,  $M_0 = M + L$ , and  $\mu_0 = \mu + l$ . Hence, we deduce that (20) is weaker than condition (9). Note also that  $r^* \leq t^*$ . Let us now show how to use method (3) until we reach some finite N such that  $x_N = y_0$  will guarantee the convergence of method (2) to a solution  $x^*$  of equation F(x) = 0. We assume that (9) does not hold for  $y_0 = x_0$  but method (3) converges to  $x^*$  (i.e., e.g. that (20) holds true). Let us assume further that

$$(43) Lr^* + l < 1.$$

Let  $z \in \bar{U}(x_0, r^*)$ . We will rewrite conditions (4)–(7), with z replacing  $x_0$ . Indeed by (7) and

(44) 
$$||A(x_0)^{-1}[A(z) - A(x_0)]|| \le L ||x_0 - z|| + l$$

$$\le Lr^* + l < 1,$$

we deduce by the Banach Lemma on invertible operators [5] that  $A(z)^{-1} \in L(Y, X)$  and

(45) 
$$||A(z)^{-1}A(x_0)|| \leq \frac{1}{1-l-L||x_0-z||} \leq \frac{1}{1-l-Lr^*}.$$

We can have using (5) and (45)

(46) 
$$||A(z)^{-1}[F'(x) - F'(y)]|| \le ||A(z)^{-1}A(x_0)|| ||A(x_0)^{-1}[F'(x) - F'(y)]||$$

$$\le \frac{K}{1 - l - L ||z - x_0||} ||x - y||$$

That is, we can set

$$(47) K_z = \frac{K}{1 - l - lr^*}.$$

Similarly in (6) we can have

(48) 
$$||A(z)^{-1}[F'(x) - A(x)]|| \le M_z ||x - x_0|| + \mu_z,$$

where,

$$M_z = \frac{M}{1 - l - Lr^*},$$

and

$$\mu_z = \frac{\mu}{1 - l - Lr^*}.$$

In view of (5), (6) and (45) for  $x \in \bar{U}$   $(x_0, r^*)$ .

(51) 
$$||A(z)^{-1}[A(x) - A(z)]||$$

$$\leq ||A(z)^{-1}A(x_0)|| [||A(x_0)^{-1}[A(x) - F'(x)]|| + ||A(x_0)^{-1}[F'(x) - F'(z)]||]$$

$$\leq \frac{1}{1 - l - lr^*} [M ||x - x_0|| + \mu + K ||x - z|| + M ||x_0 - z|| + \mu]$$

$$\leq L_z ||x - z|| + l_z,$$

where,

$$(52) L_z = \frac{K}{1 - l - Lr^*}$$

and

(53) 
$$l_z = \frac{2(Mr^* + \mu)}{1 - l - Lr^*}.$$

Moreover, in view of the estimate

(54) 
$$||A(z)^{-1}F(z)|| \le ||A(z)^{-1}A(x_0)|| \cdot ||A(x_0)^{-1}F(z)||,$$

for  $z = x_N$  we have by (32)

(55) 
$$||A(z)^{-1}F(z)|| \le \frac{1}{1 - l - Lr^*} q^N \eta,$$

where,

(56) 
$$0 \le q = \left(\frac{K}{2} + M_0\right)r^* + \mu_0 < 1,$$

That is, we can set

(57) 
$$\eta_z = \frac{1}{1 - l - Lr^*} q^N \eta.$$

Define also  $b_z$  by

$$(58) b_z = \mu_z + l_z.$$

Estimate (9) will now hold for sufficiently large N since  $\lim_{n\to\infty}q^n=0$ . Indeed, we must have

(59) 
$$\frac{2\max\{K_z, M_z\}q^N\eta}{1 - l - Lr^*} \le [1 - b_z]^2$$

which is true for

$$(60) N = \left[\frac{\ln a}{\ln q}\right] + 1$$

where [s] denotes the integer part of real number s, and

(61) 
$$a = \frac{(1 - b_z)^2 (1 - l - Lr^*)}{2 \max\{K_z, M_z\}\eta}.$$

(d) Condition (19) can be removed as follows: Define scalar sequence  $\{s_n\}$  by

(62) 
$$s_{n+2} = s_{n+1} + \frac{K}{2}(s_{n+1} - s_n)^2 + (M_0 s_n + \mu_0)(s_{n+1} - s_n), \ s_0 = 0, s_1 = \eta.$$

As in [1], [2] we can easily show that sequence  $\{s_n\}$  is nondecreasing and bounded above by

$$\bar{s} = \frac{2\eta}{2 - \delta}$$

provided that there exists  $\delta \in [0, 2)$  such that:

(64) 
$$K + \left(\frac{4M_0}{2-\delta}\right)\eta + 2\mu_0 \le \delta.$$

It follows from the proof of Theorem 2.2 that (62), (64) and  $s^* = \lim_{n \to \infty} s_n \leq \bar{s}$  can replace (19), (20),  $r^*$  respectively. Therefore we have a condition like (64) instead of (20) not requiring (19). Moreover we have

$$(65) 0 \le s_{n+1} - s_n \le \left(\frac{s}{2}\right)^n \eta.$$

We noticed that (64) may be weaker that (20). However if (64) is replaced by the stronger

(66) 
$$K\eta[1+\frac{4}{2-\delta}]+2\mu_0 \le \delta$$

or equivalently

(67) 
$$K\eta \le \frac{(\delta - 2\mu_0)(2 - \delta)}{6 - \delta}$$

then it is simple algebra to show that the right hand of (67) is always smaller than the right hand side of (20) which makes (20) weaker than condition (66) (i.e., condition (67)).

(e) It follows from (31) that condition (5) is not really needed. Indeed using only (18) and  $M_0$  instead of K in (32) condition (20) can be replaced by

(68) 
$$h'_A = M_0 \eta \le \frac{1}{2} (1 - \mu_0)^2,$$

whereas (19) is not needed. Note that in this case K is replaced by  $M_0$  in (22)–(24). If  $M_0 = K$  conditions (20) and (68) coincide. In case  $M_0 < K$  condition (68) is weaker than (20). Otherwise (20) is weaker than (68). Constant K can also be replaced by  $M_0$  in condition (64).

Let us now provide a simple numerical example in the interesting case when

(69) 
$$A(x) = F'(x) \quad (x \in D)$$

to show that (68) holds but not (9).

**Example 2.4.** Let  $X = X = \mathbf{R}$ ,  $y_0 = 1$ , D = [p, 2 - p],  $p \in [0, \frac{1}{2})$ , and define F on D by

$$(70) F(d) = d^3 - p.$$

Using (4)-(8), and (70) we obtain

(71) 
$$K = 2(2-p), \eta = \frac{1}{3}(1-p), L = 3-p, M = \mu = l = 0, \text{ and } \sigma = K.$$

Hypothesis (9) is violated since

(72) 
$$h = \frac{4}{3}(1-p)(2-p) > 1 \quad \text{for all} \quad p \in [0, \frac{1}{2}).$$

That is, there is no guarantee that method (2) converges to  $x^* = \sqrt[3]{p}$ .

In view of (18) we obtain  $M_0 = 3 - p$  and  $\mu_0 = 0$ . Condition (68) becomes

(73) 
$$h'_A = \frac{1}{3}(1-p)(3-p) \le \frac{1}{2},$$

which holds for all  $p \in \left[\frac{4-\sqrt{10}}{2}, \frac{1}{2}\right)$ .

Let us now find N such that if we set  $y_0 = z = x_N$  Newton-like method (2) will converge to  $x^*$ .

Choose p = .45. Using (20), (23) and (35), (47) and (61) we obtain

$$h_A'=.4675<\frac{1}{2}, r^*=.292176088, q=.745049024,$$
 
$$M_0=10.0019229, \bar{\eta}=1.8\bar{3} \ \ {\rm and} \ \ a=.069518717.$$

In view of (60) we obtain

$$(74) N = [9.059162824] + 1 = 10.$$

Clearly the choice of N given by (60) is in general very pessimistic. In particular in the case of Example 2.4 we have that  $x_1 = .81\overline{6}$ . Notice that

$$||F'(x_1)^{-1}F(x_1)|| = \eta_1 = .047315932.$$

It can easily be seen that (18) hold for  $x_1$  replacing  $x_0$  provided that

(75) 
$$\mu_0 = 0$$
, and  $M_0 = 6$ .

Condition (68) becomes

(76) 
$$h'_A = M_0 \eta_1 = 6(.0473)5932) = .283895592 < .5.$$

It follows from (76) that method (2) converges by simply starting from  $y_0 = x_1$  instead of  $y_0 = x_{10}$ .

#### References

- Argyros, I.K.: A unifying local-semilocal convergence analysis and applications for two-point Newton-like methods in Banach space. J. Math. Anal. Appl. 298 (2004), no. 2, 374-397.
- 2. Argyros, I.K.: Approximate solution of operator equations with applications. World Scientific Publ. Co. Hackensack New Jersey, 2005, USA.
- 3. Dennis, J.E., Jr.: On Newton-like methods. Numer. Math. 11 (1968), 324-330.
- 4. Deuflhard, P. & Heindl, G.: Affine invariant convergence theorems for Newton's method and extensions to related methods. SIAM J. Numer. Anal. 16 (1979), 1-10.
- Kantorovich, L.V. & Akilov, G.P.: Functional Analysis in Normed Spaces. Pergamon Press, Oxford, 1982.
- Ortega, J.M. & Rheinboldt, W.C.: Iterative solution of nonlinear equations in several variables. Academic Press, New York, 1970.
- Yamanoto, T.: A convergence theorem for Newton-like methods in Banach spaces. Numer. Math. 51 (1987), 545-557.

DEPARTMENT OF MATHEMATICAL SCIENCES, CAMERON UNIVERSITY, LAWTON, OK 73505, U.S.A. *Email address*: iargyros@cameron.edu